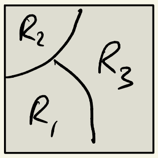
Bayes Formula

B.F. allows to swap the order of conditioning:

$$P(F|E) = \frac{P(E\cap F)}{P(E)} = \boxed{\frac{P(E|F)\,P(F)}{P(E)}} \underset{\text{L.T.P.}}{=} \frac{P(E|F)\,P(F)}{P(E|F)\,P(F) + P(E|F^c)\,P(F^c)}$$

Ex A hiker went missing in the wilderness.
She is one of the three regions with prob's 0.5, 0.3, 0.2.
Whenever anyone is lost in region 1,
the search is successful with prob. 0.4,
and for regions 2, 3 it is 0.7 and 0.8.
A search in region 1 was unsuccessful.
What is the prob. that the person is in region 1?

$S=$ 

R_i = "the hiker is in region i", $i=1,2,3$

U_i = "the search in region i is unsuccessful"

$$P(R_1|U_1) = \frac{P(U_1|R_1)\,P(R_1)}{P(U_1)} \underset{\text{L.T.P.}}{=} \frac{P(U_1|R_1)\,P(R_1)}{P(U_1|R_1)P(R_1) + P(U_1|R_2)P(R_2) + P(U_1|R_3)\,P(R_3)}$$

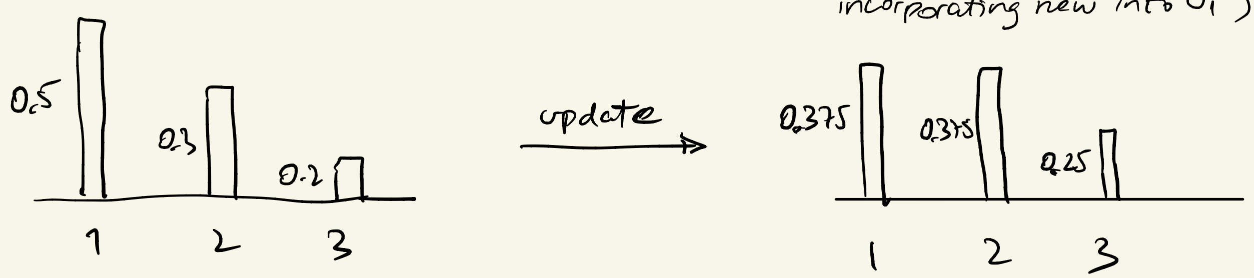
$$= \frac{(1-0.4)\,0.5}{(1-0.4)\,0.5 + 1\times 0.3 + 1\times 0.2} = 0.375$$

- Similarly, $P(R_2|U_1) = \dfrac{P(U_1|R_2)P(R_2)}{\text{same}} = 0.375$
  $P(R_3|U_1) = \dfrac{P(U_1|R_3)P(R_3)}{\text{same}} = 0.25$ } (Check!

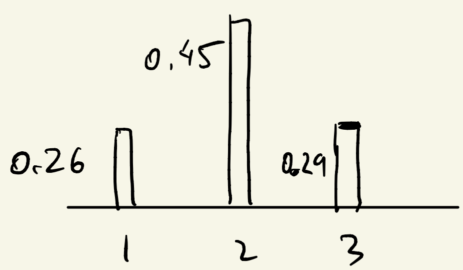
- Thus, the <u>prior</u> probabilities $P(R_1), P(R_2), P(R_3)$ got updated to the <u>posterior</u> probabilities $P(R_1|U_1), P(R_2|U_1), P(R_3|U_1)$.

<u>PRIOR</u>
(initial degree of belief)

<u>POSTERIOR</u>
(degree of belief after incorporating new info U_1)



<u>Ex</u> Another search in reg. 1 is unsuccessful
$\Rightarrow$ prob's are further updated to



(Check!)

<u>Remark</u> Bayesian model of <u>learning</u>

(text → image generative AI, spam filtering, etc.)

Def Let X be a r.v. and F be an event.
The <u>conditional expectation</u> of X given F is

$$\mathbb{E}[X|F] := \frac{\mathbb{E}[X \mathbb{1}_F]}{\mathbb{P}(F)}$$

- <u>Ex</u>: X = lifespan; F: smoker $\Rightarrow$ $\mathbb{E}[X|F]$ = life expectancy of a smoker.
- This def. is more general than the previous one:

$$\mathbb{P}(E|F) := \mathbb{E}[\mathbb{1}_E | F].$$

So we will work with <u>conditional expectation</u> from now on.

More generally:

2. Conditioning on a σ-algebra

Def Let X be a r.v, and $\mathcal{F} \subset \Sigma$ be a σ-algebra.
The <u>conditional expectation of X given $\mathcal{F}$</u> is a r.v. Y satisfying

(i) Y is $\mathcal{F}$-measurable $\leftarrow$ i.e. $\{Y \in B\} \in \mathcal{F} \ \forall B \in \mathcal{B}$

(ii) $\mathbb{E}[Y \mathbb{1}_F] = \mathbb{E}[X \mathbb{1}_F] \quad \forall F \in \mathcal{F}$

Notation: $Y = \mathbb{E}[X|\mathcal{F}]$.

- NB: $\mathbb{E}[X|\mathcal{F}]$ is a <u>random variable</u>

<u>Examples</u>:

(a) Let $\Omega = \underbrace{F_1 \sqcup F_2}_{\text{mble}}$; $\quad \mathcal{F} := \sigma(F_1, F_2) = \{\emptyset, F_1, F_2, \Omega\}$

e.g. "smoker"

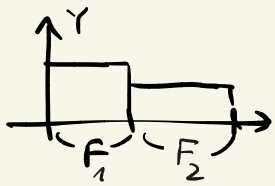
$$\Rightarrow \quad \mathbb{E}[X|\mathcal{F}] = \begin{cases} \mathbb{E}[X|F_1] & \text{if } F_1 \text{ occurs} \\ \mathbb{E}[X|F_2] & \text{if } F_2 \text{ occurs} \end{cases}$$

$\leftarrow$ life expectancy of smokers
$\leftarrow$ life expectancy of nonsmokers

$Y :=$

<u>Proof</u> (i) Y is constant on F_1 and on F_2

$\Rightarrow \forall B \in \mathcal{B}$, $\{Y \in B\}$ is $\emptyset, F_1, F_2$ or Ω $\Rightarrow \in \mathcal{F}$.



(ii) $Y \mathbb{1}_{F_1} = \begin{cases} \mathbb{E}[X|F_1] & \text{if } F_1 \text{ occurs} \\ 0 & \text{if not} \end{cases}$

$$\Rightarrow \mathbb{E}[Y \mathbb{1}_{F_1}] = \underbrace{\mathbb{E}[X|F_1]}_{\overset{\text{def}}{=} \frac{\mathbb{E}[X \mathbb{1}_F]}{\mathbb{P}(F)}} \cdot \mathbb{P}(F_1) = \mathbb{E}[X \mathbb{1}_{F_1}].$$

Similarly for $F_2, \emptyset, \Omega$. □

(b) More generally, consider a partition

$$\Omega = F_1 \sqcup \cdots \sqcup F_n$$

and let $\mathcal{F} := \sigma(F_1, \dots, F_n)$. Then

$$\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X \mid F_i] \text{ if } F_i \text{ occurs.} \qquad (\text{Check!})$$

e.g. $F_1 = \{\text{person's age} \in [0,10)\}$, $F_2 = \{\text{age} \in [10,20)\}$, ..., $F_{10} = \{\text{age} \in [90,100)\}$

X = person's height

$\Rightarrow \mathbb{E}[X \mid \mathcal{F}]$ = ave. height of the random person's cohort

(c) Same example, in the analytic form:

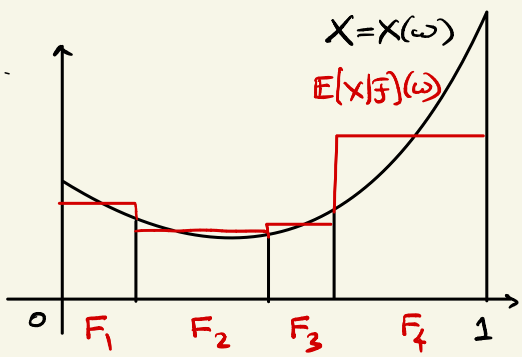
$\Omega = [0,1]$, $\Sigma = \mathcal{B}(\mathbb{R})$, P = Lebesgue measure.

$X : [0,1] \to \mathbb{R}$ an integrable function.

$$\mathbb{E} X = \int_0^1 X(\omega)\, d\omega$$

$$\mathbb{E}[X \mid F_i] = \frac{1}{|F_i|} \int_{F_i} X(\omega)\, d\omega$$

$$\mathbb{E}[X \mid \mathcal{F}](t) = \frac{1}{|F_i|} \int_{F_i} X(\omega)\, d\omega \quad \text{if } t \in F_i$$



Remarks

(a) For the <u>coarsest</u> σ-algebra $\mathcal{F} = \{\emptyset, \Omega\}$, $\quad \mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}(X)$

For the <u>finest</u> σ-algebra $\mathcal{F} = \Sigma$, $\quad \mathbb{E}[X \mid \Sigma] = X$

Intermediate $\mathcal{F}$ $\Rightarrow$ interpolates between $\mathbb{E}[X]$ and X.

(b) $\mathcal{F}$ encodes <u>available information</u> (e.g. smoking habits, age, ...)

$\mathbb{E}[X \mid \mathcal{F}]$ encodes the best prediction of X given that info

No info $\Rightarrow \mathbb{E}[X]$ $\qquad$ All info $\Rightarrow X$ (exact)

3. Existence, Uniqueness

THM $\forall$ integrable r.v. X and $\forall$ σ-algebra $\mathcal{F} \subset \Sigma$,

(i) $\mathbb{E}[X|\mathcal{F}]$ exists

(ii) and is unique: if Y, Y' are both cond. exp's of X given $\mathcal{F}$, then $Y = Y$ a.s.

Proof of Existence is based on:

THM (Radon-Nikodym Thm)

Let μ, ν be two σ-finite measures on a measurable space $(\Omega, \mathcal{F})$.

(Ω = countable union of mble sets with finite measures)

Assume $\nu \ll \mu$ (absolutely continuous)

($\mu(F)=0 \Rightarrow \nu(F)=0 \ \forall F \in \mathcal{F}$)

Then $\exists$ $\mathcal{F}$-measurable function $f: \Omega \to [0,\infty)$

s.t. $\nu(F) = \int_F f \, d\mu \quad \forall \ F \in \mathcal{F}$

$:= \int f \mathbb{1}_F \, d\lambda$

- WLOG $X \geq 0$ (otherwise decompose $X = X^+ - X^-$)
- Define $\nu(F) := \mathbb{E}[X \mathbb{1}_F]$, $F \in \mathcal{F}$.
- Then ν is a finite measure on $(\Omega, \mathcal{F})$

since $\mathbb{E}X < \infty$

$\forall$ disjoint $F_1, F_2, \ldots \in \mathcal{F}$:

$$\nu\left(\bigcup_{i=1}^\infty F_i\right) = \mathbb{E}\left[\sum_{i=1}^\infty X \mathbb{1}_{F_i}\right] \overset{\text{by monotone convergence thm}}{=} \sum_1^\infty \mathbb{E}[X \mathbb{1}_{F_i}] = \sum_1^\infty \nu(F_i)$$

- $\nu \ll \mathbb{P}$ (if $\mathbb{P}(F)=0$ then $\mathbb{E}[X \mathbb{1}_F] = 0$)
- Apply RNT $\Rightarrow$ $\exists$ $f =: Y$ that is $\mathcal{F}$-integrable and

$$\mathbb{E}[X \mathbb{1}_F] = \int_F Y \, d\mathbb{P} = \mathbb{E}[Y \mathbb{1}_F] \quad \forall \ F \in \mathcal{F}. \qquad \square$$

Proof of Uniqueness Assume $\mathbb{E}[X \mathbb{1}_F] = \mathbb{E}[Y \mathbb{1}_F] = \mathbb{E}[Y' \mathbb{1}_F]$ $\forall F \in \mathcal{F}$

Subtract $\Rightarrow$ $\mathbb{E}[(Y-Y')\mathbb{1}_F] = 0$. Apply for $F := \{Y - Y' \geq 0\}$ $\Rightarrow$

$\mathbb{E}(Y-Y')^+ = 0$. Similarly, $\mathbb{E}(Y-Y')^- = 0$.

Add $\Rightarrow$ $\mathbb{E}|Y-Y'| = 0$ $\Rightarrow$ $Y = Y'$ a.s. $\square$